DOCUMENT RESUME

ED 330 717 TM 016 304

AUTHOR	Yates, Beverly J.
TITLE	A Comparison of Effectiveness Ratings of Selected Principals and NASSP Assessment Center Ratings.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	Apr 91
NOTE	24p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 3-7, 1991).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Administrator Evaluation; *Assessment Centers (Personnel); Case Studies; Comparative Analysis; Instructional Leadership; Interrater Reliability; Interviews; *Predictive Validity; *Principals; Secondary Education; Test Validity
IDENTIFIERS	*Diagnostic Assessment School Principal Effect; National Association Secondary School Principals

ABSTRACT

The predictive validity of the National Association of Secondary School Principals (NASSP) assessment center evaluation process for principals is compared with the perceived effectiveness of a selected population of principals. The NASSP assessment center approach includes a case study, a personal interview, two exercises, and a scholastic examination. The tasks are assigned to address problem analysis, judgment, organizational ability, decisiveness, leadership, sensitivity, stress tolerance, oral communication, written communication, personal motivation, range of interests, and educational values. The present correlational study used scores from 44 principals who had attended two midwestern assessment centers. To obtain ratings of principal effectiveness, the selected principals were then asked to complete the principal form of the Diagnostic Assessment of School and Principal Effectiveness (DSAPE). The Pearson correlation coefficient was used to measure the association between the sets of scores. With 312 correlations computed for the selected sets, it would be expected that chance alone would reveal 15 significant correlations at the 0.05 level. In this study, six correlations were judged to be significant at the 0.05 level, and none were judged to be significant at the 0.01 level. There seems to be no relation between DSAPE and NASSP assessment center scores. Ten tables and a 27-item list of references are included. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
BEVERLY J. YATES
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED330717

A Comparison of Effectiveness Ratings of Selected Principals and NASSP Assessment Center Ratings

Beverly J. Yates
University of Kansas

BEST COPY AVAILABLE

April 1991

Paper presented at the annual meeting of the American Educational Research Association, Chicago, 1991. Funding for this study was provided by KanLEAD, the U.S. Department of Education LEAD Program 84-178.

TM014304

I. Introduction

The current era in education is being defined in terms of reform, criticism, and initiatives and mandates on the national, state, and local levels. Emphasis is being placed on school improvement with little agreement existing on what facet of education can or should be altered in order to result in the final product required to sustain or remake our social order. Each segment of society holds a somewhat different view of the role of education and therefore defines the problem, or lack thereof, slightly differently, resulting in a wide spectrum of proposed solutions, desired outcomes, methodologies, research directions, and degree of commitment to any of the above.

However, the role of the educational administrator, particularly that of the principal, has been increasingly identified as a pivotal one. The principal is the most visible representation of schools to large segments of the community, and the role of the principal is being focused on as the one most important in the provision of effective schools. One of the emphases of these change efforts has become the identification of principals who have the skills and abilities to meet the demands of society for education and provide leadership for the effective delivery of education.

Assessment center evaluation is one of the tools being employed by educators in the effort to select administrators in a rational way according to defined standards. The assessment center process is a set of exercises or tests designed to provide evaluation of an individual's strengths and weaknesses with reference to a particular position. These exercises are designed to simulate tasks which a person holding that position would need to be able to execute to successfully work in that position. The exercises are completed in a controlled environment and observed and rated by a group of trained assessors. The resulting outcome provides a detailed description of the participant's performance on those specific exercises. Most often applicants or aspirants to a managerial position are nominated for an assessment center when an organization seeks to predict probable administrative performance.

As the emphasis on effective schools has increased, the attendant concern with improvement of methods of principal selection has grown, and with it an increased use of educational assessment centers. One of the most prevalently used systems was developed by The National Association of Secondary School Principals (NASSP). Assessment center evaluation is costly in comparison to other methods of selection. This cost may fall to the district, state, educational organization, or candidate. Whether the candidate is required to make this investment in time and dollars or a public agency accepts the responsibility, the process should be significantly more effective than less expensive methods of selection to justify this use of resources.

The value of an assessment center as a selection tool is dependent on the level of content and criterion-related validity which can be established. A validity study of the NASSP system was conducted at Michigan State University (Schmitt, Noe, Meritt, Fitzgerald, & Jorgensen, 1984) with the cooperation of NASSP. Overall correlations between performance ratings and assessment center ratings were positive (.21 to .30). However, correlations of support staff ratings of principal performance and assessment center ratings were lower (.05). It is the only published study relating to the validity of the NASSP process.

The purpose of this study was to compare the predictive usefulness of the NASSP assessment center process with the perceived effectiveness of a selected population of principals and, thereby, add to the body of literature concerning the selection process.

II. Review of Literature

The assessment center concept of employee selection was first utilized on a large scale by German military psychologists prior to World War II (Bray & Grant, 1966). The next documented use of the method is by the United States Office of Strategic Services in selection of operatives during World War II (Moses, 1977a). Current assessment center techniques, used to identify potential managers or administrators, are constructed on the same model and the objective of assessment continues to be "prediction of behavior as based on critical examination of behavior" (MacKinnon, 1977).

After World War II, use of assessment centers generally disappeared until the 1950s (Moses, 1977a). The Management Progress Study (Bray, 1964) was initiated by American Telephone & Telegraph (AT&T) in 1956, revitalizing the use of the assessment center process. The purpose of this study was "to learn about the characteristics and growth of men as they become, or try to become, the middle and upper managers of a large concern" (Bray, 1964, p. 420). As an individual became a part of the study, information was collected through a wide range of techniques. Included was an assessment center procedure. Follow-up studies and evaluations were done.

The AT&T study is basic to the development of assessment centers for several reasons. It was a longitudinal study and was ongoing until the time that the courts required AT&T to divest itself of the regional telephone companies in the 1980s. The result is almost 30 years of data collection and analysis. The study was kept separate from operational procedures, and information on individuals was not released. As a result of the preservation of confidentiality, criterion-related validity studies of the various procedures employed could be completed without contamination of assessment results (Bray & Grant, 1966). The subjects were reassessed so that the effects of growth on the assessed characteristics could be measured.

Many other Fortune 500 companies adopted the assessment center method and expanded its use from evaluation of supervisory potential to middle and higher management assessment, management trainee selection, sales potential assessment, and identification of engineering skills (Byham, 1977; Hinrichs, 1878; Mitchel, 1975). An example in the public sector is the use of an assessment center to select administrative interns in Philadelphia (Strausbaugh & Wagman, 1977). According to Byham (1977), educational use was limited to the university level in evaluation of students for graduate programs.

Byham (1970) included employee self-development and career planning as useful outcomes of an assessment center process. He also recommended the application of the results to design jobs which match a particular person's abilities and potential. Hersey (1982) and Nickerson (1986) suggested that techniques and materials from assessment centers be used in graduate programs or in counseling graduate students in educational administration.

Validity Studies

Bray and Grant established the predictive validity of the assessment center (Sackett, 1982) with correlations of .44 and .71 between assessment center prediction and management level attained for the college and non-college samples.

Cohen, Moses, and Byham (cited in Sackett, 1982) summarized 18 validity studies and reported an average correlation of .40 with managers' ratings and number of promotions received, and an average correlation of .63 with managers' ratings of a candidate's promotion potential. However, criterion contamination may have affected the outcomes as results of the assessments were available to managers except in the two AT&T studies cited above.

History of the NASSP Assessment Center Process

In 1974, a pilot project known as the Technical Assistance Program was developed by the American Psychological Association's Division of Industrial and Organizational Psychology (Moses, 1977b). The objective was to offer professional services to an agency in the public sector. NASSP was identified as a nationally based organization that had considerable influence in a major public area, and was able to generalize results of a project to other systems. The assessment center method was selected as the project and the American Psychological Association provided technical assistance (Hersey, 1977). Jeswald was the lead technical advisor and described the process of designing an assessment center (Jeswald, 1977).

The tasks in an NASSP assessment center evaluation include a case study, a personal interview, two in-basket exercises, and a scholastic exam (Schmitt, Meritt, Fitzgerald, & Noe, 1982). The tasks are designed to assess 12 dimensions: problem analysis, judgment, organizational ability, decisiveness, leadership, sensitivity, stress tolerance, oral communication, written

communication, personal motivation, range of interests, and educational values (Schmitt et al., 1984).

In 1975, the Prince William County and Charlottesville school districts in Virginia were selected to participate in the pilot project (Hersey, 1977). NASSP provided methodology, materials, and assessor training. Psychologists from the American Psychological Association assisted in monitoring the activities in order to protect the integrity of the psychometric process. Four years were spent in research and development (Hersey, 1986). Since that time the number of NASSP centers has increased to about 60 and the primary use has been evaluation of prospective principals.

Value of the Assessment Center Process

The advantages of using an assessment center method for employee selection varies with the institution. The criterion-related validities of the AT&T study have demonstrated the predictive value of the procedure (Bray & Grant, 1966). The Michigan State study of the NASSP system (Schmitt et al., 1982) established a positive relationship between assessment center ratings and performance evaluations. Reilly and Chao (1982) reviewed validity studies of eight selection procedures: biodata, interviews, peer evaluations, self assessments, reference checks, academic performance, expert judgments, and projective techniques. Assessment center ratings were included in standardized tests which had criterion-related validities of .35 to .45.

In addition to the obvious usefulness of a technique which may increase effectiveness in selection, the implementation of the 1978 Uniform Guidelines on Employee Selection Procedures has increased concern that employers be able to clarify and evaluate selection procedures. If the assessment center has been developed using job analysis procedures to define the exercises (Olshfski & Cunningham, 1986), content validity can generally be established. The Equal Employment Opportunity Commission has generally viewed assessment centers favorably, although some researchers express concerns. Sackett (1987), for example, suggests that since administrators commonly go through supervisory training programs after being selected, the measurement during selection of skills and knowledge requisite to the position is not content valid.

Most criticisms of the assessment center method have developed in the private sector and are related to validity issues. Hinrichs (1978) found that manager evaluations of potential had a higher correlation (.55) to management level achieved after eight years than did the assessment center ratings (.46). He justified the use of the assessment center on the basis of providing data to clarify and evaluate the promotion system.

Sackett (1987) criticized content validity practices in assessment center studies, as referenced above. He proposed the use of criterion-related validity. Norton (1977, 1981) debated this issue with Dreher and Sackett (1981) in a series of articles.

III. Research Methodology

This research is a correlational study of assessment center ratings and effectiveness ratings of selected principals. Scores were obtained from two NASSP assessment centers. To obtain ratings of principal effectiveness the selected principals were then invited to take part in the research by completing the Principal form of the Diagnostic Assessment of School and Principal Effectiveness (DASPE) (Ebmeier, 1989). Their supervisors were invited to complete the Supervisor form and five of each principal's staff were invited to complete the Staff form.

Identification of the Population

The population to be studied was defined as persons who had completed a NASSP assessment center, subsequently acceded to the principalship, and held that position for one school year. All persons from two midwestern NASSP assessment centers who met those criteria were invited to participate in the study. There was a total of 44 possible participants. A decision was made to treat the entire population rather than to sample the small number. Further, the descriptive data for the population would be stratified by gender, race, and building level, rather than treating small numbers of cases in a stratified population.

Data Collection

Data from the two assessment centers was collected for each of the persons eligible for the study. The overall score, as well as the scores for each of the 12 dimensions which are rated in the NASSP process were collected.

Each of the 44 principals received packets which contained a letter inviting them to participate in the research, an informed consent form, the Principal form of the DASPE, and a response sheet. Postage was provided for the return responses. Enclosed in the principals' packets were five packets for staff members. Instructions were provided for sampling an alphabetical roster of the staff to identify participants. Staff members were provided with separate return envelopes which were coded and returned directly to the researcher. In addition, each principal's supervisor received a similar packet.

As an incentive to participate in the research and as a show of gratitude for that participation, each principal received a profile which detailed the school and individual's effectiveness as rated on the DASPE. They were also informed of the number of persons who had contributed to that profile.

For the purpose of analysis, the data collected from staff members was collapsed for each principal, providing an average score. Data for the supervisor, principal and staff was maintained separately and separate correlations with assessment center scores were calculated.

IV. Presentation of the Data

The Population

The population to be studied was defined as persons who had completed a NASSP assessment center, subsequently acceded to the principalship, and held that position for one school year. There were 44 persons identified who had attended either of two midwestern NASSP assessment centers. All 44 were invited to participate in the study, as well as their supervisors and five of their staff members. Some level of response was received from 42 of the 44 units. Of the principals, 24 returned completed Principal forms of the DASPE. The response rate for supervisors was 22. At least one staff member responded for 37 of the principals. A complete response set (the principal, supervisor and 5 staff members) was received in 4 cases. Table 1 reports the response rate for each case. In Table 2, the rate of response for sets of respondents is shown.

TABLE 1

SURVEY RESPONSE RATE BY CASE

Case	Respondent			Case	Respondent		
	S[a]	P[b]	St[c]		S[a]	P[b]	St[c]
1	X	X		23		X	3
2		X	4	24	X	X	3
3			1	25	X		2
4			4	26			4
5			2	27	X	X	3
6			3	28	X		2
7	X	X	2	29	X		
8	X			30		X	2
9		X	2	31	X		2
10	X		2	32	X	X	2
11		X	4	33	X	X	3
12		X	2	34		X	5
13	X	X	5	35		X	3
14				36		X	4
15	X			37	X		
16			4	38			2
17	X	X	5	39	X	X	3
18		X	2	40	X		1
19		X	1	41	X	X	3
20		X	4	42	X	X	5
21	X		3	43			
22			3	44	X	X	5

[a] Supervisor
[b] Principal
[c] Number of Staff

TABLE 2

SURVEY RESPONSE RATE BY SETS OF RESPONDENTS

Respondent Set	Rate	Respondent Set	Rate
S[a], P[b], St 5[c]	4	P, St 5	1
S, P, St 4	0	P, St 4	5
S, P, St 3	5	P, St 3	2
S, P, St 2	2	P, St 2	4
S, P, St 1	0	P, St 1	1
S, P	1	P	0
S, St 5	0	St 5	0
S, St 4	0	St 4	2
S, St 3	1	St 3	2
S, St 2	4	St 2	2
S, St 1	1	St 1	1
S	4	0	2

Note. N = 44.

[a] Supervisor

[b] Principal

[c] Number of Staff

When the population was identified, 44 principals met the criteria for participation in the study. It was decided to treat the population rather than develop a sample from a population of 44. No stratification occurred. Demographics for the population are reported in Table 3.

TABLE 3

DEMOGRAPHICS OF POPULATION

Descriptor		Frequency	Percent
Gender	Female	31	70.5
	Male	13	29.5
Race	Black	6	13.6
	Hispanic	2	4.5
	White	36	81.8
Building Level	Elementary	29	65.9
	Middle	6	13.6
	High	9	20.5

Note. N = 44.

Outcomes

The measure of association between the sets of scores was the Pearson correlation coefficient. Correlations were computed for the NASSP overall and dimension scores with each of the dimensions of effectiveness available from the DASPE, resulting in 312 combinations.

While the NASSP data were complete for each subject, the DASPE data were directly affected by the response rate (See Tables 1 & 2). There were 7 possible respondents for each principal, the principal, supervisor and 5 staff members, and each respondent may or may not have elected to participate in the study. Therefore, the number of respondents represented by the DASPE data is variable. The set of data containing all supervisors and principals was used to calculate correlations for those respondents. A series of computations were completed with varying degrees of the population to ascertain the effect of the number of staff respondents for the principals. The set of computations for staff results reported for this study included the 22 cases for which at least 3 staff members reported. Correlations were incomplete in sets with fewer cases. Computations with sets which involved smaller numbers of respondents per case resulted in widely varying results.

With 312 correlations computed for the selected sets, it would be expected that there could be 15 significant correlations at the .05 level by chance. In this study 6 were judged to be significant at the .05 level and 0 at the .01 level. There is not sufficient statistical evidence to support the original hypothesis of significant correlations between the two sets of scores.

Additionally, it should be noted that 191 correlations did not reach the .2000 level. Of the remaining 121 which exceeded .2000, 38 exceeded .3000, and 12 were in excess of .4000. Analysis of the significance of the Pearson product moment correlations (Cohen, 1977) results in a power of .23 for correlations of .20 with a population of 22 and a power of .40 for correlations of .30. For statistically significant results a power of .80 would be desirable. For correlations of .20 to have a power of .80, 200 cases would be required. Eighty-four cases would be needed to result in a power of .80 for correlations of .30.

Supervisors' reports of effectiveness more frequently resulted in agreement with assessment center ratings with 54 of the 121 correlations exceeding .2000 being associated with supervisor scores. Principals created the next largest group with 36 and the staff set was smallest with 31 of the 121.

The assessment center dimension showing the largest number of correlations exceeding .2000 was judgment at 18. The remaining dimensions in descending order were: written communication, 16; decisiveness, 13; organizational ability, 11; oral communication, 10; problem analysis, 9; personal motivation and educational values, 8; range of interests and leadership, 6; stress tolerance, 5; and sensitivity, 4. There were 7 correlations exceeding .2000 associated with the overall score.

The statistical results for the Pearson correlation coefficient are tabulated in Tables 13 to 19. The 191 correlations <.2000 are not reported.

TABLE 4

CORRELATIONS OF PRINCIPAL RESPONSES TO DASPE REGARDING THE SCHOOL WITH PRINCIPAL SCORES FROM NASSP ASSESSMENT CENTER

	DASPE			
NASSP	Maintenance	Adaptation	Goal Attainment	Integration
Overall				.2095 p=.326
Problem Analysis	.3017 p=.152			.3387 p=.105
Judgment			.2971 p=.159	
Organizational Abil.	.3816 p=.066	.2787 p=.187		.4380* p=.032
Decisiveness			.3727 p=.073	.2779 p=.188
Leadership		.2571 p=.255		
Sensitivity	.2637 p=.213	.3748 p=.071	.2143 p=.315	.2957 p=.161
Range of Interests			.2182 p=.306	
Personal Motivation		.2329 p=.273		
Educational Values				
Stress Tolerance				
Oral Communication				
Written Communication	.3825 p=.065	.4174* p=.042		.3500 p=.094

Note. N = 24. Correlations <.2000 were not reported.

*p <.05.

TABLE 5

CORRELATIONS OF PRINCIPAL RESPONSES TO DASPE REGARDING THEIR OWN BEHAVIOR WITH PRINCIPAL SCORES FROM NASSP ASSESSMENT CENTER

NASSP	DASPE			
	Maintenance	Adaptation	Goal Attainment	Integration
Overall		.2441 p=.286		
Problem Analysis		.3728 p=.096		
Judgment	.3431 p=.109	.5016* p=.021	.2674 p=.217	.4425* p=.034
Organizational Ability				
Decisiveness		.2065 p=.369		
Leadership		.2352 p=.305		
Sensitivity				
Range of Interests	.3025 p=.161	.2772 p=.224	.2536 p=.243	
Personal Motivation				
Educational Values		.2927 p=.198		
Stress Tolerance		.3735 p=.095		.2084 p=.340
Oral Communication		.2460 p=.282		.2532 p=.244
Written Communication				

Note. N = 21 for Adaptation and 23 for Maintenance, Goal Attainment, and Integration. Correlations <.2000 were not reported.

*p <.05.

TABLE 6

CORRELATIONS OF SUPERVISOR RESPONSES TO DASPE REGARDING THE SCHOOL WITH PRINCIPAL SCORES FROM NASSP ASSESSMENT CENTER

	DASPE			
NASSP	Maintenance	Adaptation	Goal Attainment	Integration
Overall				
Problem Analysis	.2086 p=.352	.2070 p=.355	.2461 p=.270	
Judgment	.2121 p=.343	.2947 p=.183	.3317 p=.132	.3383 p=.124
Organizational Ability		.2162 p=.334	.2831 p=.202	.2109 p=.346
Decisiveness		.2099 p=.349	.2891 p=.192	.2172 p=.332
Leadership				.2559 p=.250
Sensitivity				
Range of Interests	.2115 p=.345			
Personal Motivation	.2499 p=.262	.2455 p=.271	.2008 p=.370	.2242 p=.316
Educational Values	.4534* p=.034	.3729 p=.087	.3530 p=.107	.2474 p=.267
Stress Tolerance	.2719 p=.221	.3216 p=.144		
Oral Communication		.2850 p=.199	.2088 p=.351	.2586 p=.245
Written Communication		.2994 p=.176	.2772 p=.212	.2743 p=.217

Note. N = 22. Correlations <.2000 were not reported.

*p <.05.

TABLE 7

CORRELATIONS OF SUPERVISOR RESPONSES TO DASPE REGARDING PRINCIPAL BEHAVIORS WITH PRINCIPAL SCORES FROM NASSP ASSESSMENT CENTER

	DASPE			
NASSP	Maintenance	Adaptation	Goal Attainment	Integration
Overall		.2164 p=.404	.3339 p=.190	
Problem Analysis			.2941 p=.252	
Judgment		.2673 p=.300	.3423 p=.179	.2870 p=.264
Organizational Ability		.2157 p=.406	.3643 p=.151	.2159 p=.405
Decisiveness				
Leadership		.2933 p=.253	.2828 p=.271	.2924 p=.255
Sensitivity				
Range of Interests				
Personal Motivation			.2167 p=.404	.2239 p=.388
Educational Values	.4361 p=.080	.4113 p=.101	.4799 p=.051	
Stress Tolerance				
Oral Communication	.4188 p=.094	.4136 p=.099	.5366* p=.026	.3984 p=.113
Written Communication	.2196 p=.399		.3315 p=.194	

Note. N = 17. Correlations <.2000 were not reported.

*p <.05.

TABLE 8

CORRELATIONS OF STAFF RESPONSES TO DASPE

REGARDING THE SCHOOL WITH PRINCIPAL SCORES FROM

NASSP ASSESSMENT CENTER

	DASPE			
NASSP	Maintenance	Adaptation	Goal Attainment	Integration
Overall				
Problem Analysis		.2209 p=.323		
Judgment	.2698 p=.225	.3696 p=.090	.2263 p=.311	.2520 p=.258
Organizational Abil.	.2821 p=.203	.2785 p=.209		
Decisiveness	.2723 p=.220	.4298* p=.046	.3194 p=.147	
Leadership				
Sensitivity				
Range of Interests				
Personal Motivation	.2595 p=.244			
Educational Values				
Stress Tolerance		.2103 p=.347		
Oral Communication		.3036 p=.170		
Written Communication	.2692 p=.226	.3557 p=.104	.2777 p=.211	.2486 p=.209

Note. The average staff response was calculated for each principal. The set of cases reported included 22 cases for which at least 3 staff members reported. Correlations <.2000 were not reported.

*p <.05.

TABLE 9

CORRELATIONS OF STAFF RESPONSES TO DASPE

REGARDING PRINCIPAL BEHAVIORS WITH PRINCIPAL SCORES FROM

NASSP ASSESSMENT CENTER

	DASPE				NASSP
	Mainte-nance	Adapta-tion	Goal Attainment	Inte-gration	
Overall			.2156 p=.335	.2140 p=.339	.2152 p=.336
Problem Analysis					
Judgment			.2343 p=.294	.2114 p=.345	
Organizational Ability					
Decisiveness		.2639 p=.235	.2734 p=.218	.3042 p=.169	.2816 p=.204
Leadership					
Sensitivity					
Range of Interests			.2203 p=.325		
Personal Motivation					
Educational Values					
Stress Tolerance					
Oral Communication					
Written Communication		.2087 p=.351	.2722 p=.220	.2298 p=.304	.2426 p=.277

Note. The average staff response was calculated for each principal. The set of cases reported includes the 22 cases for which at least 3 staff members reported. Correlations <.2000 were not reported.

*p <.05.

TABLE 10

SUMMARY OF DISTRIBUTION OF CORRELATIONS

Correlations		Rate
<.2000		191
>.2000	Supervisors	54
	Principals	36
	Staff	31
	Total	121
	NASSP Dimension	
	Judgment	18
	Written Communication	16
	Decisiveness	13
	Organizational Ability	11
	Oral Communication	10
	Problem Analysis	9
	Personal Motivation	8
	Educational Values	8
	Overall	7
	Range of Interests	6
	Leadership	6
	Stress Tolerance	5
	Sensitivity	4

Note. Total correlations computed = 312. Significant at .05 level = 6.

V. Implications of This Study

Results were obtained from two NASSP assessment centers. The overall score, as well as the scores for each of the 12 dimensions were collected. The population was defined as principals who had participated in a NASSP assessment center, acceded to the principalship and held that position for one year. There were 44 persons identified who met those criteria.

It was possible to receive data from the DASPE on each of the 44 subject principals from seven sources: (1) the principal, (2) the supervisor, (3) staff member #1, (4) staff member #2, (5) staff member #3, (6) staff member #4, (7) staff member #5. Of the 44 subject principals, some feedback was received from at least one source in 42 cases. The responses were as follows: principals, 24; supervisors, 22; and at least one staff member for 37 cases.

Therefore, the number of respondents represented by the data is variable. The set of data containing all supervisors and principals was used to calculate correlations for those respondents. A series of computations were completed with varying degrees of the population to ascertain the effect of the number of staff respondents for the principals. The set of computations for staff results reported for this study included the 22 cases for which at least 3 staff members reported. Correlations were incomplete in sets with fewer cases. Computations with sets which involved smaller numbers of respondents per case resulted in widely varying results.

The measure of association between the sets of scores was the Pearson correlation coefficient. Correlations were computed for the NASSP overall and dimension scores with each of the dimensions of effectiveness available from the DASPE, resulting in 312 combinations.

With 312 correlations computed for the selected sets, it would be expected that there could be 15 significant correlations at the .05 level by chance. In this study 6 were judged to be significant at the .05 level and 0 at the .01 level. There is not sufficient statistical evidence to support the original hypothesis of significant correlations between the two sets of scores.

Additionally, it should be noted that 191 correlations did not reach the .2000 level. Of the remaining 121 which exceeded .2000, 38 exceeded .3000, and 12 were in excess of .4000. Analysis of the significance of the Pearson product moment correlations (Cohen, 1977) results in a power of .23 for correlations of .20 with a population of 22 and a power of .40 for correlations of .30. For statistically significant results a power of .80 would be desirable. For correlations of .20 to have a power of .80, 200 cases would be required. Eighty-four cases would be needed to result in a power of .80 for correlations of .30. As a result of the low power of these results, no inferences should be drawn.

The original validity study of the NASSP system conducted at Michigan State University reported that overall correlations

between performance ratings and assessment center ratings were positive (.21 to .30). The sample size for NASSP data was 425. Performance data of at least one type was collected on 167 principals. There were 110 cases in the sample which resulted in the correlation of .21 for the overall score with self ratings and 121 for the overall correlation of .30 for teachers. Supervisor (n = 120) ratings resulted in a correlation of .25.

The relatively small number of cases in this study and the low power of results do not entirely account for the lack of significant correlations. It should be noted that the number of correlations at the .05 level of significance was 6 when 15 could have been expected by chance. There does not seem to be a relationship between scores resulting from ratings of principal effectiveness and NASSP assessment center scores. The NASSP center may not address the same dimensions as are identified as behaviors found in effective principals. This study is not strong enough to allow a judgment but does raise a question about selecting a principal on the basis of high NASSP assessment center performance when the objective is to identify a person who will become a strong principal as described by effective schools literature..

When a district chooses a selection tool the issues of validity of the instrument, time required in administration and cost must be addressed. Validities of selection tools as reported by Reilly and Chao (1982) are: biodata, .35; interviews, .19; peer evaluations, .41; self assessment, .15; reference checks, .14; academic performance, .17; expert judgment, .17; projective techniques, .18; and standardized tests, .35 - 45. Assessment centers were included with standardized tests.

As cited above, the Michigan study of the NASSP process produced validities in the .21 - .30 range. Districts commonly use interviews, reference checks, academic performance, and expert judgments. Each of these methods of selection result in lower validities than assessment centers, including the NASSP process. They are also low cost.

Districts, especially in states where an assessment center is required for certification, may not have to assume the cost or invest the time required to do the assessing. When that is the case, the NASSP process falls in the same cost range as the other commonly used instruments and has slightly higher validity. Districts may then determine that the assessment center is a useful source of information, in combination with other tools, in formulating a decision about principalship candidates.

If a district is to assume the cost of the assessment center or the cost of a candidate's assessment, different factors must be reviewed. The cost then becomes many times higher than the other selection tools while the validity remains somewhat higher. If the district is providing the assessors through the use of district administrators, the cost in time away from normal duties must also be considered.

Another use of the NASSP assessment center process is to determine certification at the state level. The correlations derived in the Michigan study are not exceptional. There are no

published studies addressing the issue of prediction of performance as an effective principal. It is an expensive investment for states or for candidates. The careers of candidates are advanced or constrained on the basis of assessment center outcomes. Given this, it is particularly important that the use of the NASSP assessment center process as a determiner of certification should be examined carefully.

Limitations of the Study

One significant limitation to this study was the size of the population. There were 44 persons identified with 42 participating. Variation in the response rate further decreased the number of cases in the study. In order to increase the population size, additional NASSP assessment centers should be involved.

Of major concern to this study was the effect of restriction of range. Candidates to the NASSP assessment centers involved in the study are generally nominated by their school districts and may not represent the total population of persons who would be candidates for the principalship. The range could be further restricted by the effect of a poor assessment center score. Persons who do not do well at the assessment center may not be offered the opportunity to accede to the principalship. Lastly, scores are derived in the NASSP assessment center process through consensus. This eliminates scores which would deviate significantly.

As a condition of access to scores from one of the two NASSP assessment centers, the researcher did not have direct contact with over 50% of the population. This was a limiting factor to the response rate as follow-up was difficult.

Recommendations for Further Study

In order to increase the size of the potential population, additional NASSP assessment centers should be involved. Principals are interested in participating in a study which provides them with a profile describing their effectiveness and could be contacted through assessment center records. It is indicated that more research be conducted concerning the relationship between the NASSP assessment center process and effective schools research.

References

Bray, D. W. (1964). The management progress study. American Psychologist, 19, 419-429.

Bray, D. W. & Grant, D. L. (1966). The assessment center in measurement of potential for business management. Psychological Monographs, 80 (17, Whole No. 625).

Byham, W. C. (1970). Assessment centers for spotting future managers. Harvard Business Review, 48, 150-167.

Byham, W. C. (1977). Application of the assessment center method. In J. L. Moses & W. C. Byham, (Eds.), Applying the Assessment Center Method, (pp. 3-11). New York: Pergamon Press.

Cohen, J. (1977). Statistical Power Analysis for the Behavioral Sciences (Rev. ed.). New York: Academic Press.

Dreher, G. & Sackett, P. (1981). Some problems with applying content validity evidence to assessment center procedures. Academy of Management Review, 6, 551-560.

Ebmeier, H, (1989, June). The development and field test of an instrument for client based principal formative evaluation (1,2). Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Hersey, P. W. (1977). NASSP's assessment center--From concept to practice. NASSP Bulletin, 61, 74-76.

Hersey, P. W. (1982). NASSP assessment center develops leadership talent. Educational Leadership, 39, 370-371.

Hersey, P. W. (Ed.) (1986). Selecting and developing educational leaders: A search for excellence. NASSP's assessment center project. NASSP Bulletin, 70, 1-2.

Hinrichs, J. K. (1978). An eight-year follow-up of a management assessment center. Journal of Applied Psychology, 63, 596-601.

Jeswald, T. A. (1977). A new approach to identifying administrative talent. NASSP Bulletin, 61, 79-83.

MacKinnon, D. W. (1977). From selecting spies to selecting managers--the OSS assessment program. In J. L. Moses & W. C. Byham, (Eds.), Applying the Assessment Center Method (pp. 13-30). New York: Pergamon Press.

Mitchel, J. O. (1975). Assessment center validity: A longitudinal study. Journal of Applied Psychology, 60, 573-579.

Moses, J. L. (1977a). The assessment center method. In J. L. Moses & W. C. Byham, (Eds.), Applying the Assessment Center Method (pp. 3-11). New York: Pergamon Press.

Moses, J. L. (1977b). Developing an assessment center program for school administrators. NASSP Bulletin, 61, 76-79.

Nickerson, N. (1986). Affecting preservice, inservice programs at universities. In P. W. Hersey (Ed.), Selecting and developing educational leaders: A search for excellence. NASSP Bulletin, 70, 56-58.

Norton, S. (1977). The empirical and content validity of assessment centers vs. traditional methods for predicting managerial success. Academy of Management Review, 2, 442-453.

Norton, S. (1981). The assessment center process and content validity; A reply to Dreher and Sackett. Academy of Management Review, 6, 561-566.

Olshfski, D. F. & Cunningham, R. B. (1986). Establishing assessment center validity; An examination of methodological and theoretical issues. Public Personnel Management, 15(1), 85-98.

Reilly, R. R. & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. Personnel Psychology, 35, 1-62.

Rist, M. C. (1986). Principals mull the merits of new evaluation techniques. Executive Educator, 8(4), 37,43.

Sackett, P. R. (1982). A critical look at some common beliefs about assessment centers. Public Personnel Management, 2, 140-147.

Sackett, P. R. (1987). Assessment centers and content validity; Some neglected issues. Personnel Psychology, 40(1), 13-25.

Schmitt, N., Meritt, R., Fitzgerald, M. P., & Noe, R. A. (1982). The NASSP assessment center; a validity report. NASSP Bulletin, 66, 134-42.

Schmitt, N., Noe, R., Meritt, R., Fitzgerald, M., & Jorgensen, C. (1984). Criterion-related and Content Validity of the NASSP Assessment Center. Michigan State University, Department of Psychology.

Strausbaugh, D. & Wagman, G. (1977). An assessment center examination to select administrative interns. Public Personnel Management, 6, 263-269.